NASA Center for Computational Sciences (NCCS)
Greenbelt, MD

Attendees:

Nancy Palm   Head, Science Computing Branch

Bruce Pfaff   System Administration

Joe Schmidt   Applications Support (CSC)

http://nccsinfo.gsfc.nasa.gov/scb

"The overarching purpose of Earth system science is to develop the knowledge basis for predicting future changes in the coupled physical, chemical, geological, biological, and social state of the Earth and assessing the risks associated with such change."

NASA Earth Science Research Strategy
for 2000-2010, p. 17
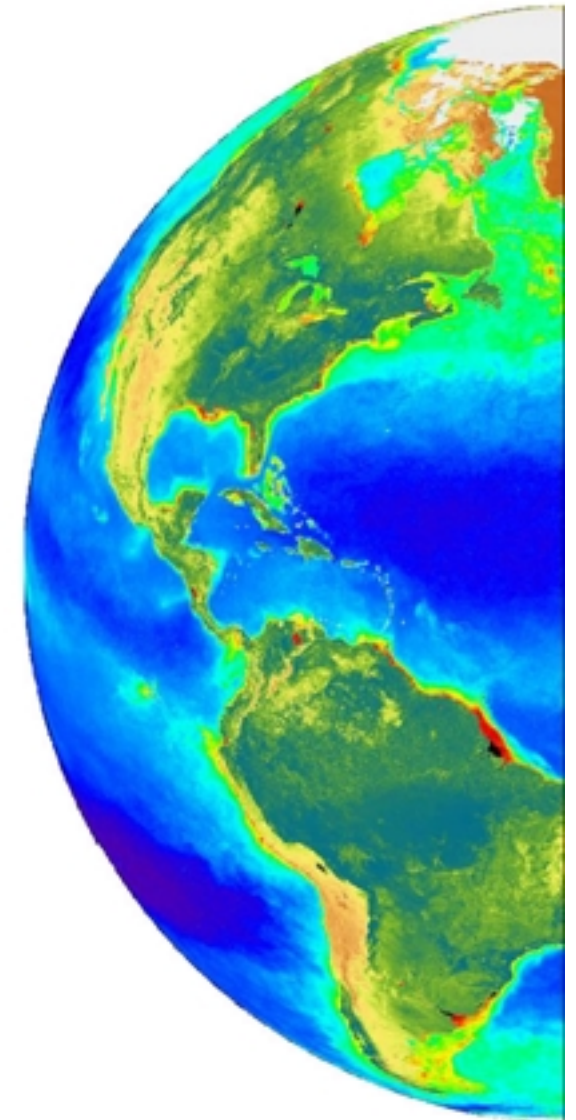
NCCS User Community

- ## NASA Seasonal-to-Interannual Prediction Project
  - Ensemble simulations coupling high-resolution atmospheric and oceanic models to understand near-term climate phenomena such as El Niño-La Niña.
  - Migrating from a Cray T3E.
  - Moderate CPU count, short duration runs which require scheduling flexibility such as pre-emptive scheduling for high priority jobs.

- ## Data Assimilation Office
  - Creates research-quality assimilated global data sets from multiple satellites for climate analysis and observation planning.
  - Migrating from SGI Origins.
  - Requires stable environment with high-availability.
  - Operational requirement to generate a forecast product every 6 hours.

September 23, 2002

NCCS User Community (Cont.)

- **Goddard Institute for Space Studies**
  - Climate studies focusing on timescales ranging from a decade to a century, specializing in exploring the natural and human influences on global warming.
  - Migrating from SGI Origins.
  - Low CPU count, short-duration runs.

- **ESTO/Computational Technologies Project**
  - Research computational technology for the advancement of Earth and space science.
  - Developing the Earth System Modeling Framework (ESMF).
  - Migrating from a diverse collection of platforms.
  - Requires occasional access to the full machine.

September 23, 2002

NCCS User Community (Cont.)

- **Research and Analysis Groups**
  - A large collection of smaller research efforts which individually have modest computing requirements. Examples include:
    - Studying large-scale tropical circulations.
    - Developing and testing a cloud liquid water and microphysics parameterization.
    - Investigating the predictability of the North American summer climate.
    - Modeling the magnetic fields from the Earth's core, lithosphere, ionosphere and magnetosphere.
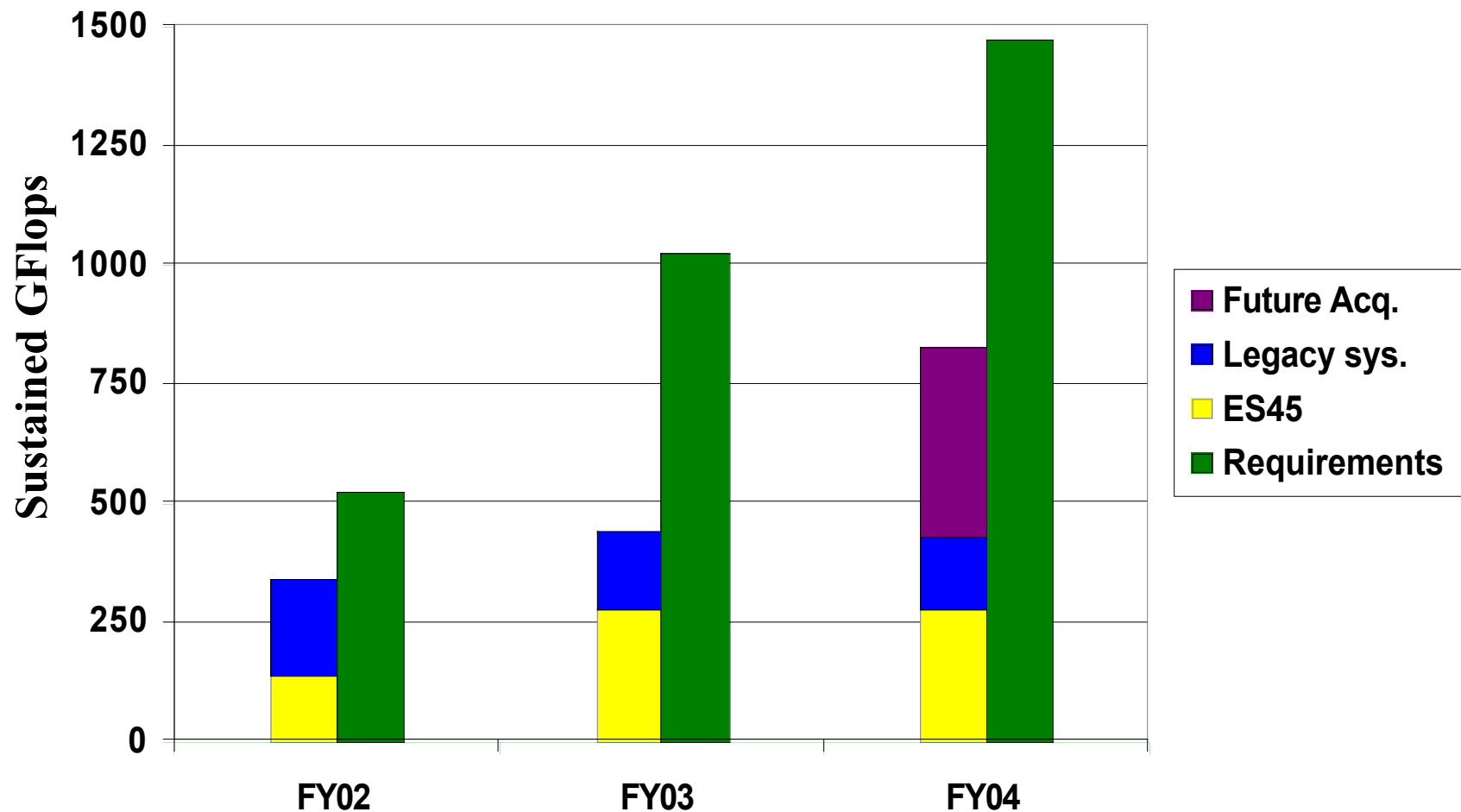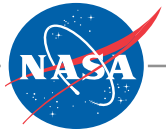 - Migrating primarily from Cray SV1 and SGI Origin.

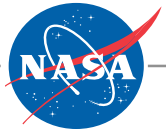# Computing Requirements

September 23, 2002
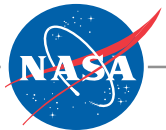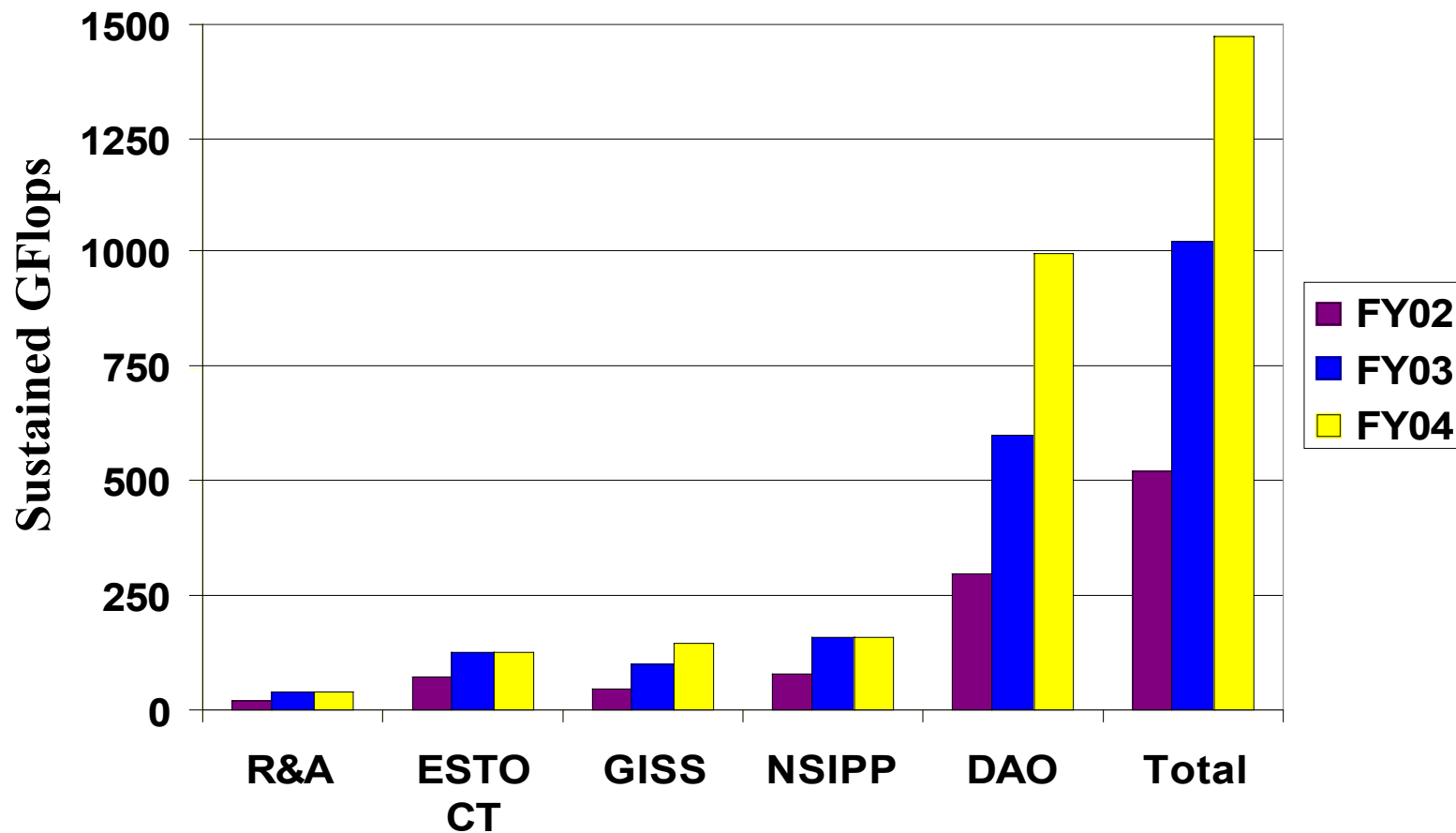
Requirements and Resources

The Acquisition Methodology

- To meet the exploding computational requirements, the NCCS has radically altered its acquisition process.

- Merged previously separate funding streams to gain the advantage of economies-of-scale for Earth science production computing.

- Two-year cycle for solicitations of new systems maintains state-of-the-art capabilities to take advantage of Moore's law.

- Use of integrator (CSC)

September 23, 2002

NCCS Computational Capability

System	CPUs	Peak Gflops	Memory (GBs)	Disk (TBs)
HP/ Compaq ES45	1392	3200	440	8.5
HP/ Compaq ES45	32	64	13	0.5
SGI Origin 3000	512	409.6	128	3.5
SGI Origin 2000	64	38.4	32	11.5
Cray T3E	1360	778	162	2.2
Cray SV1	32	32	8	1.2

September 23, 2002

NASA Center for Computational Sciences
Total Data Stored Original 10/1999 Requirements

Legend:
- DAO Risk Mitigation
- DAO
- NSIPP Risk Mitigation
- NSIPP
- Carbon Cycle Risk Mitigation
- Carbon Cycle
- Research & Analysis Risk Mitigation
- Research & Analysis
- GISS Risk Mitigation
- GISS
- ESTO/CT Risk Mitigation
- ESTO/CT
- Pre-FY2000 Risk Mit.
- Pre-FY2000

ems   9/19/2002

Projections from "Refined Earth Science Computing Requirements" 10/21/1999
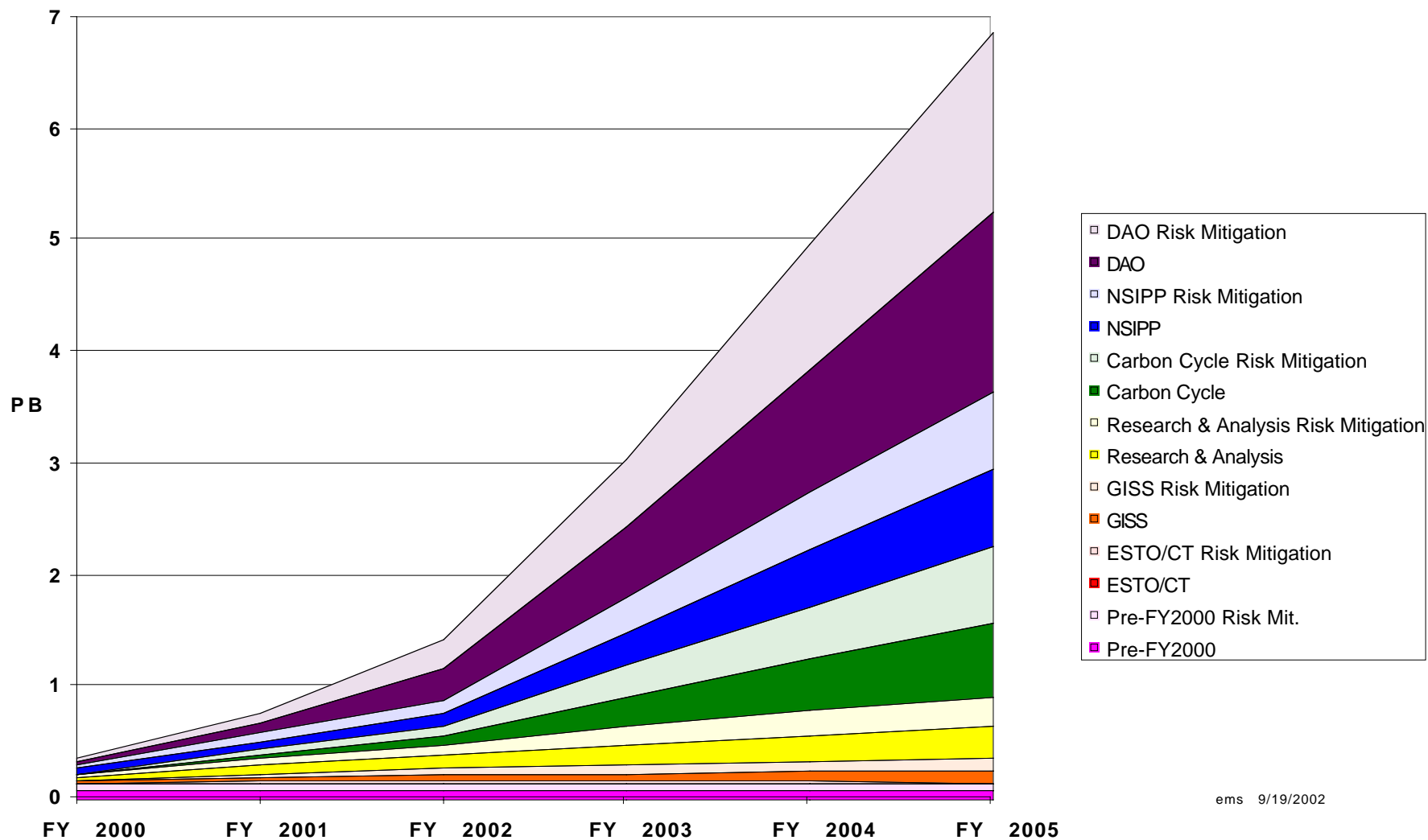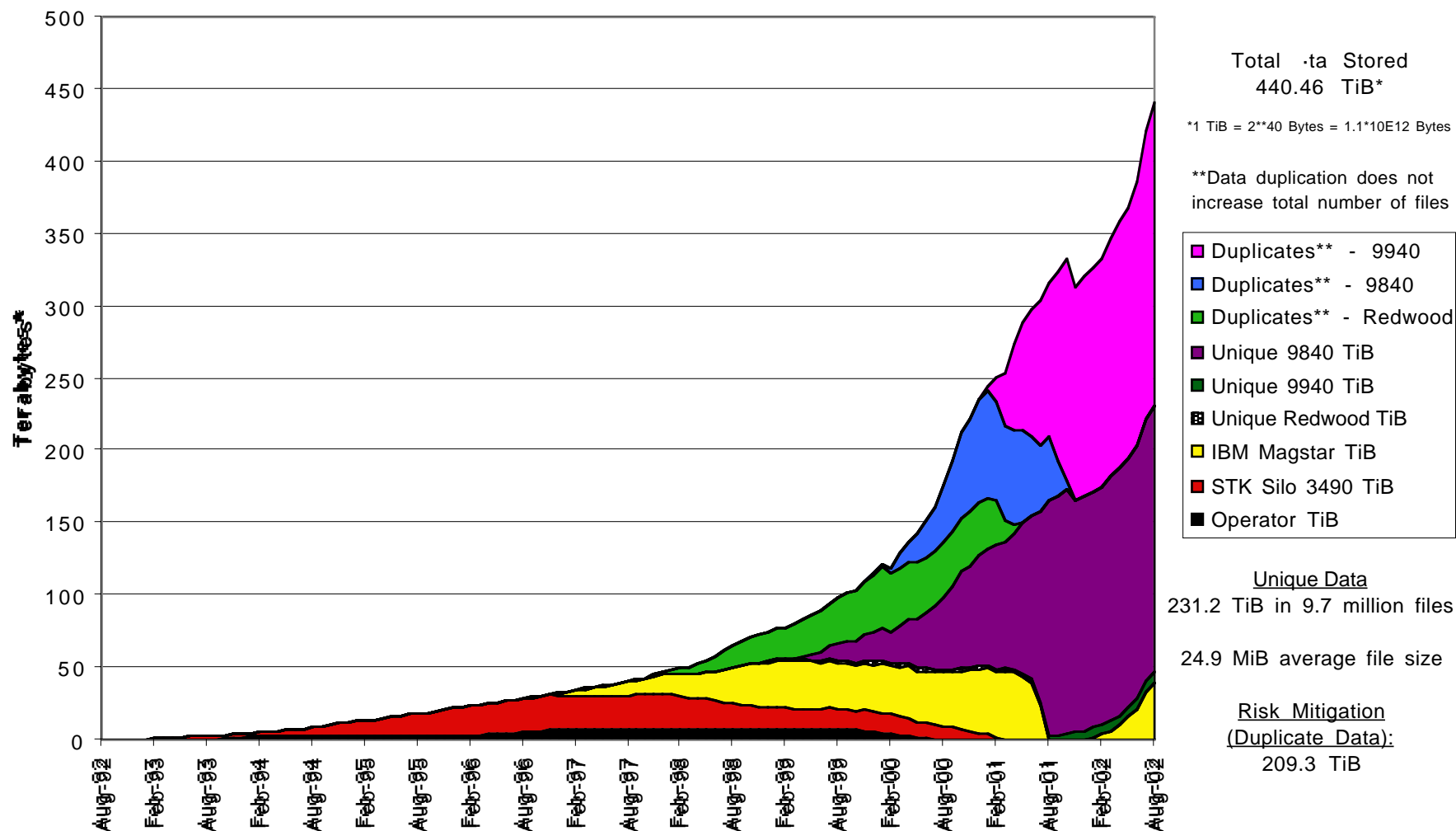
ems 3/24/2000

September 23, 2002

10

NASA Center for Computational Sciences
Mass Data Storage and Delivery System
Total Terabytes Stored

Total ·ta Stored
440.46 TiB*

*1 TiB = 2**40 Bytes = 1.1*10E12 Bytes

**Data duplication does not increase total number of files

- ■ Duplicates** - 9940
- ■ Duplicates** - 9840
- ■ Duplicates** - Redwood
- ■ Unique 9840 TiB
- ■ Unique 9940 TiB
- ▦ Unique Redwood TiB
- ■ IBM Magstar TiB
- ■ STK Silo 3490 TiB
- ■ Operator TiB

Unique Data
231.2 TiB in 9.7 million files

24.9 MiB average file size

Risk Mitigation
(Duplicate Data):
209.3 TiB

ems 9/19/2002

11

September 23, 2002

HP Challenges/Difficulties

- Compiler problems with use of advanced F90 features.
- Difficulty maintaining concurrent versions of system software (compilers and libraries) as required by the exacting validation process of our users.
- RMS/LSF integration has a myriad of problems which have terminated simulations that depend on a stable batch system. Optimistic about Kite release.
- Machine robustness - the inability of the current version of the software to handle gracefully problems with head nodes.
- Many other problems have been identified and addressed in a timely manner by the HP team.
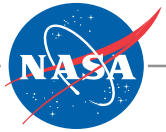
Future Challenges

- Scheduling for five major user groups with different requirements and nodes with two clock speeds.
- Providing capability for application level checkpoint/restart initiated by the system.
- Facilitating code and data migration.
- Achieving "lights out" operation.
- Increasing efficiency of machine utilization.
- Managing large volumes of data.
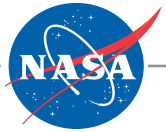
Conclusions

- We have achieved 90% utilization on the initial 512 processor installment.

- 1392 processor system is scheduled to be available this week!

- The NCCS wants to work with the HP community to benefit from shared experiences.

- Hardware alone cannot meet the computational challenge the NCCS faces. The challenge must be met by improvements in software.